

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 13, 2003, 01:25:43 ; Search time 3327 Seconds  
(without alignments)  
7480.548 Million cell updates/sec

Title: US-09-864-921-97  
Perfect score: 5277  
Sequence: 1 MNFIKDSRALIQRMGMTVI.....WQFDDDLVSITGAFKLVT A 1024

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DB=frame\_plus\_p2n.model -DEV=xlh  
-Q/cgn2\_1/USPTO.spool/US9864921/runat\_11122003\_092025\_9940/app query.fasta 1.1223  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US9864921 @CGN 1 1 2477 @runat\_11122003\_092025\_9940 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hcc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hcc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_nam.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_fod.\*  
26: em\_gss\_phg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3496	66.2	3224	11 AK035300	AK035300 Mus muscu
2	952	18.0	552	28 AQ309404	AQ309404 CITBI-EI-
3	934	17.7	697	28 BH293386	BH293386 CH230-44G
4	838	15.9	650	10 BB627584	BB627584 BB627584
5	815	15.4	741	12 BI824482	BI824482 603038854
6	801	15.2	480	9 AV719179	AV719179 AV719179
7	799	15.1	602	28 AZ360053	AZ360053 IM0103H11
8	777	14.7	748	12 BI908869	BI908869 603066455
9	693	13.1	518	28 BH348412	BH348412 CH230-42F
10	605.5	11.5	619	28 BH267158	BH267158 CH230-19B
11	599	11.4	371	9 AV656315	AV656315 AV656315
12	585	11.1	364	9 AI263294	AI263294 qx57b01.x
13	585	11.1	775	12 BI854236	BI854236 60381263
14	564.5	10.7	817	28 BH358172	BH358172 CH230-18E
15	563	10.7	503	13 BX109471	BX109471 BX109471
16	561	10.6	670	14 BY732706	BY732706 BY732706
17	524	9.9	637	9 AL639997	AL639997 AL639997
18	514	9.7	719	13 BQ204082	BQ204082 UI-R-DN1-
19	475.5	9.0	476	10 BG210375	BG210375 RST29913
20	472	8.9	728	9 AL652549	AL652549 AL652549
21	451.5	8.6	470	28 AQ624020	AQ624020 HS 5378 B
22	436.5	8.3	423	14 CB547126	CB547126 AMGNNUC.TS
23	430	8.1	636	13 BU076330	BU076330 im49ph08.x
24	416	7.9	645	29 BZ759604	BZ759604 622 IL4 C
25	409	7.8	637	9 AL782121	AL782121 AL782121
26	377	7.1	261	9 AW337918	AW337918 hel2h11.x
27	375	7.1	482	29 BZ760724	BZ760724 623 4L4 H
28	323	6.1	404	28 AQ889169	AQ889169 HS 2161_B
29	303	5.7	1058	12 BM477340	BM477340 AGENCOURT
30	298	5.6	553	28 AQ320928	AQ320928 RPCI11-93
31	292	5.5	478	13 BY558831	BY558831 BY558831
32	268	5.1	4065	11 AK089662	AK089662 Mus muscu
33	262	5.0	466	13 BY557516	BY557516 BY557516
34	260	4.9	219	28 AQ283886	AQ283886 RPCI11-78
35	260	4.9	427	13 BZ111342	BZ111342 BZ111342
36	257	4.9	577	12 BJ037401	BJ037401 BJ037401
37	255	4.8	630	28 AQ112439	AQ112439 CIT-HSP-2
38	253	4.8	4442	11 AK089843	AK089843 Mus muscu
39	250	4.7	714	9 AL657756	AL657756 AL657756
40	245	4.6	3218	11 AK054378	AK054378 Mus muscu
41	243	4.6	3577	11 AK087843	AK087843 Mus muscu
42	237.5	4.5	251	10 BF903662	BF903662 IL2-MT018
43	234	4.4	3368	11 BC021272	BC021272 Homo sapi
44	234	4.4	3369	11 BC012789	BC012789 Homo sapi
45	232	4.4	625	28 AZ614134	AZ614134 IM0442N17

ALIGNMENTS

RESULT 1  
LOCUS AK035300  
DEFINITION Mus musculus adult male urinary bladder cdna, RIKEN full-length enriched library, clone:953001i19 product:similar to CASPASE RECRUITMENT DOMAIN PROTEIN 12 (ICE-PROTEASE ACTIVATING FACTOR) (IPAF) (CARD, LRR, AND NACHT-CONTAINING PROTEIN) (CLAN PROTEIN) [Homo sapiens], full insert sequence.  
ACCESSION AK035300  
VERSION AK035300.1  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

3224 bp mRNA linear HTC 05-DEC-2002

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

## AUTHORS

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

## MEDLINE

## PUBMED

99279253

10349636

## REFERENCE

## AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

## MEDLINE

## PUBMED

20499374

11042159

## REFERENCE

## AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system-384-format

sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

## MEDLINE

## PUBMED

20530913

11076861

## REFERENCE

## AUTHORS

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,

Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,

Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,

Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,

Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,

Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,

Kuehl, P., Lewis, S., Matsumoto, Y., Nikaide, I., Pesole, G.,

Quackenbush, J., Schriml, L. M., Stauble, F., Suzuki, R., Tomita, M.,

Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,

Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,

Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,

Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,

Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,

Marchionni, L., Mashima, J., Mazzarelli, J., Mombærte, P., Nordone, P.,

Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,

Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,

Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,

Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.,

and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 685-690 (2001)

## MEDLINE

## PUBMED

21085660

11217851

## REFERENCE

## AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

## MEDLINE

## PUBMED

1217851

## REFERENCE

## AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,

Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,

Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,

Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,

Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,

Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,

Okazaki, Y., Saito, R., Saichoh, H., Sakai, C., Sakai, K., Sakazume, N.,

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,

Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,

Muramatsu, M., and Hayashizaki, Y.

Direct Submission

## JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/.

## FEATURES

## source

Location/Qualifiers  
1. .3224  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:9530011P19"  
/db\_xref="taxon:10090"  
/clones="9530011P19"  
/sex="male"  
/tissue\_type="urinary bladder"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
<1. .2693  
/note="putative"  
similar to CASPASE RECRUITMENT DOMAIN PROTEIN 12  
(ICE-PROTEASE ACTIVATING FACTOR) (IPAF) (CARD, LRR, AND  
NACHT-CONTAINING PROTEIN) (CLAN PROTEIN) [Homo sapiens]  
(SWISSPROT|Q9NPP4, evidence: FASTA, 76%ID, 87.5%length,  
match=2689)"  
3205. .3210  
polyA\_signal  
3224  
polyA\_site  
3224  
/note="putative"  
BASE COUNT 929 a 754 c 779 g 762 t  
ORIGIN

## misc\_feature

/note="putative"  
similar to CASPASE RECRUITMENT DOMAIN PROTEIN 12  
(ICE-PROTEASE ACTIVATING FACTOR) (IPAF) (CARD, LRR, AND  
NACHT-CONTAINING PROTEIN) (CLAN PROTEIN) [Homo sapiens]  
(SWISSPROT|Q9NPP4, evidence: FASTA, 76%ID, 87.5%length,  
match=2689)"  
3205. .3210  
polyA\_signal  
3224  
polyA\_site  
3224  
/note="putative"

BASE COUNT 929 a 754 c 779 g 762 t

## ORIGIN

Alignment Scores:  
Pred. No.: 0 Length: 3224  
Score: 3496.00 Matches: 681  
Percent Similarity: 85.06% Conservative: 82  
Best Local Similarity: 75.92% Mismatches: 133  
Query Match: 66.25% Indels: 1  
DB: 11 Gaps: 0

US-09-864-921-97 (1-1024) x AK035300 (1-3224)

QY 129 AsnLeuTySerThrPheThrGluProIleLeuTrpArgLysAspGlnHisHisArg 148

Db 2 AATCTGGAGAAACCTTCACAGAACCTATCATCTGGAGAGAGGACCATCGTCATCCCGT 61

QY 149 ValGluGlnLeuThrLeuAsnGlyLeuGlnAlaLeuGlnSerProCysIleIleGlu 168

Db 62 GTGGAGCAGCTGACTTTTGGCGAGCCTGTCTGAGGCTCTGAAGAGCCCTCGCTGATTGAA 121

QY 169 GlyGluSerGlyLysGlyLysSerThrLeuLeuGlnArgIleAlaMetLeuTrpGlySer 188

Db 122 GCGAGTCTGCAAGAGGAGGAGTCCACCTGCTGCAGAGATCGCTATGCTCTGGGCTCT 181

QY 189 GlyLysCysLysAlaLeuThrLysPheLysPheValPhePheLeuArgLeuSerArgala 208

Db 182 GGGGGCTGCAGGGCTCTGAAGGGTTCAGATTAGTCTTCTTCATCCACCTGAGAGCGCC 241

QY 209 GlnGlyGlyLeuPheGluThrLeuGlnGlnLeuLeuAspIleProGlyThrIleArg 228

Db 242 ATGGGGGAGCTATTTCGAAGAACTGTACATCAGCTCTGCAACATACCCGCTTCATCAGC 301

QY 229 LysGlnThrPheMetAlaMetLeuLeuLysLeuArgGlnArgValLeuPheLeuLeuAsp 248

Db 302 AAGCCGACCTTCAAGGCTCTGCTGCTGAAGCTACACAAGAGGTCCTCTTTCTTCTCGAT 361  
Qy 249 GlyTyrAsnGluPheLysProGlnAsnCysProGluIleGluAlaLeuIleLysGluAsn 268  
Db 362 GGTTCATGAATTCATCCCGAGACTCCCGAGAAATGGAAGCCCTGTAAAGGAAC 421  
Qy 269 HisArgPheLysAsnMetValIleValThrThrThrGluCysLeuArgHisIleArg 288  
Db 422 CATCGCTTCAAGAACATGCTGTCATGTCACCAACCCAGGAGTCTGAGGCATATCAGA 481  
Qy 289 GlnPheGlyAlaLeuThrAlaGluValGlyAspMetThrGluAspSerAlaGlnAlaLeu 308  
Db 482 CATGTTGGCGCCCTGACTCGGAGGTGGAGATATGACCAAGACAGAGTCCCAAGATCTC 541  
Qy 309 IleArgGluValLeuIleLysGluLeuAlaGluGlyLeuLeuGlnIleGlnLysSer 328  
Db 542 ATCGAGGCGAGTCTGCTGCTGATGAGTTCAGGCTGTCGCGCCCAAAATCCAGGAGTCC 601  
Qy 329 ArgCysLeuArgAsnLeuMetLysThrProLeuPheValIleThrCysAlaIleGln 348  
Db 602 AGGTGCTGAGAAATCTGATGAAGACCCCTCTCTTCGTGCTGATCACCTGTCAATTCAG 661  
Qy 349 MetGlyGluSerGluPheHisThrHisThrGlnThrThrLeuPheHisThrPheTyrAsp 368  
Db 662 ATGGGACAGACAGGAATTCAGAGCTCACACCAACCATGCTGTCCAAACCTTCTACGAC 721  
Qy 369 LeuLeuIleGlnLysAsnLysHisLysHisLysGlyValAlaAlaSerAspPheIleArg 388  
Db 722 CTCCTGATACAGAAACAGCCACAGATATAGAGTGGAGCTTCAGGTGATTTGCCAGG 781  
Qy 389 SerLeuAspHisArgGlyAspLeuAlaLeuGluGlyValPheSerHisLysPheAspPhe 408  
Db 782 AGCCTAGACTACTGTGGAGCTGGCCCTGAGAGGTGTTGCGCCACAAATTCATTT 841  
Qy 409 GluLeuGlnAspValSerValAsnGluAspValLeuLeuThrThrGlyLeuLeuCys 428  
Db 842 GAACCCGAGCATGGGTCCAGCATGAACGAGGACGCTCTGTCACAAATAGGGCTCCTCTGT 901  
Qy 429 LysTyrThrAlaGlnArgPheLysProLysTyrLysPhePheHisLysSerPheGlnGlu 448  
Db 902 AAGTACACAGCTCAGAGGCTGAAGCCCGACGTATAAATCTTTCAATAATCATTTCAGGAG 961  
Qy 449 TyrThrAlaGlyArgArgLeuSerSerLeuLeuThrSerHisGluProGluGluValThr 468  
Db 962 TACACGGCAGGTGGAGACTCAGCAGTTGCTGACGTCGTCAGAGCCAGAGAGGTGAGC 1021  
Qy 469 LysGlyAsnGlyTyrLeuGlnLysMetValSerIleSerAspIleThrThrTyrSer 488  
Db 1022 AAAGGGAAACAGCTACTTAAACAAAATGGTTTCATCTCTGACATCACATCCCTATATGGC 1081  
Qy 489 SerLeuLeuArgTyrThrCysGlySerSerValGluAlaThrArgAlaValMetLysHis 508  
Db 1082 AATCTGCTCTCTACACGCTGGGTGCTCCACAGAACCAACAGGCGGTCTATGAGGCAC 1141  
Qy 509 LeuAlaAlaValTyrGlnHisGlyCysLeuLeuGlyLeuSerIleAlaLysArgProLeu 528  
Db 1142 CTTGCAATGTTTATCAGACCGCAGCCTCAGAGACTTTCAGTCACCAAGAGGCTCTC 1201  
Qy 529 TrpArgGlnGluSerLeuGlnSerValLysAsnThrThrGluGlnGluIleLeuLysAla 548  
Db 1202 TGGAGGACAGGAATCAATCAGAGTCTGAGAAATACCACTGAGCAAGATGTTCTGAAAGCC 1261  
Qy 549 IleAsnIleAsnSerPheValGluCysGlyIleHisLeuTyrGlnGluSerThrSerLys 568  
Db 1262 ATCAATGTAAATTCCTTCGTAGAGTGTGGCATCAATTTGTTCTCAGAGAGTATGTCTAAA 1321  
Qy 569 SerAlaLeuSerGlnGluPheGluAlaPhePheGlnGlyLysSerLeuTyrIleAsnSer 588  
Db 1322 TCAGACCTGAGCCAGAAATTCAGGCTTTCTTTCAAGGTPAAAGTTTATACATCACTCA 1381  
Qy 589 GlyAsnIleProAspTyrLeuPheAspPhePheGluHisLeuProAsnCysAlaSerAla 608  
Db 1382 GAGAACATCCCTGACTATTATTATTTGACTTCTTTTGAATACTTGCCTAAATTTGTCAAGGCCA 1441

Qy 609 LeuAspPheIleLysLeuAspPheTyrGlyGlyAlaMetAlaSerTrpGluLysAlaAla 628  
Db 1442 TTGGACTTCGTGAAGTTGGATTTCTATGAAGAGCTACAGAGTACAGAGCAAGCAGNA 1501  
Qy 629 GluAspThrGlyGlyIleHisMetGluGluAlaProGluThrTyrIleProSerArgAla 648  
Db 1502 GAGAATGTCCTGGAGTTTACACAGAGGGCCCTCAGAAACCTACATATCCCCCGGGCT 1561  
Qy 649 ValSerLeuPhePheAsnTrpLysGlnGluPheArgThrLeuGluValThrLeuArgAsp 668  
Db 1562 GTGCTTTGTTCTTCACTGGAAGCAGGAATTCAGAGCTCTAGAGGTACACTCCGAGAT 1621  
Qy 669 PheSerLysLeuAsnLysGlnAspIleArgTyrLeuGlyLysIlePheSerSerAlaThr 688  
Db 1622 ATTACAGTTGTAATACAGAGATATCAATATCTGGGGAAGATATTAGCTCTGCCACC 1681  
Qy 689 SerLeuArgLeuGlnIleLysArgCysAlaGlyValAlaGlySerLeuSerLeuValLeu 708  
Db 1682 AACCTCGGCTGCATATCAAGAGATGTGCAGCATGTGCGAGACTTGGAAAGACTCAGCTCAGTCTC 1741  
Qy 709 SerThrCysLysAsnIleTyrSerLeuMetValGluAlaSerProLeuThrIleGluAsp 728  
Db 1742 AGAACCCTCAAGAACATCATACCTCATGTTGGTGAAGCCAGTCCCTCACCACGGATGAC 1801  
Qy 729 GluArgHisIleThrSerValThrAsnLeuLysThrLeuSerIleHisAspLeuGlnAsn 748  
Db 1802 GAACAGTACATACATCTGTGACAGGCTCCAGAACTTAAGTATTTCACCGTTGCACACT 1861  
Qy 749 GlnArgLeuProGlyGlyLeuThrAspSerLeuGlyAsnLeuLysAsnLeuThrLysLeu 768  
Db 1862 CAACAGCTGCCAGGTGCTGATTGACAGCTTGGTATCTCTGAAGAACTCGAGAGACTC 1921  
Qy 769 IleMetAspAsnIleLysMetAsnGluGluAspAlaIleLysLeuAlaGluGlyLeuLys 788  
Db 1922 ATACTGGATGACATCAGGATGAACGAGGAAGATGCTAAAAACCTAGCCGAGGCTACGA 1981  
Qy 789 AsnLeuLysLysMetCysLeuPheHisLeuThrHisLeuSerAspIleGlyGluGlyMet 808  
Db 1982 AGCCTGAAGAGATGCGTTTACTCTCATTTGACTCATTTGTCTGACATTTGGGGGGGATG 2041  
Qy 809 AspTyrIleValLysSerLeuSerGluProCysAspLeuGluGluIleGlnLeuVal 828  
Db 2042 GACTACATAGTCAAGTCTCTCTCAGAAAGATCCTGTGATCTCCAAGAGATGAAGTTGGTG 2101  
Qy 829 SerCysCysLeuSerAlaAsnAlaValLysIleLeuAlaGlnAsnLeuHisAsnLeuVal 848  
Db 2102 GCCTGCTGCTGACTGCAAACTCTGTGAAAGTTCTAGCACAGAAATCTTCACAATTTGATC 2161  
Qy 849 LysLeuSerIleLeuAspLeuSerGluAsnTyrLeuGluLysAspGlyAsnGluAlaLeu 868  
Db 2162 AAGCTGACATCTTTGATATATCAAAAATTTACCTGGAAAGAGTGGGAATGAAGCTCTA 2221  
Qy 869 HisGluLeuIleAspArgMetAsnValLeuGluGlnLeuThrAlaLeuMetLeuProTrp 888  
Db 2222 CAGGAACATGATCGCAGGCTTGGCGTCTCTGGAGAGCTCACTACATTTGCTGCTTGG 2281  
Qy 889 GlyCysAspValGlnGlySerLeuSerSerLeuLysHisLeuGluGluValProGln 908  
Db 2282 TGCTGGGATGTGCACACCAGCTGCCAAGCTGTTGAAGCAGATTGGAGGGGACCCACGA 2341  
Qy 909 LeuValLysLeuGlyLysAsnTrpArgLeuThrAspThrGluIleArgIleLeuGly 928  
Db 2342 CTTGCCAACTTGGATTGAAAACCTGGAGACTCAGAGACGAAGATTAAGAATTTAGGT 2401  
Qy 929 AlaPhePheGlyLysAsnProLeuLysAsnPheGlnGlnLeuAsnLeuAlaGlyAsnArg 948  
Db 2402 GAATTTCTGGAGATCAATCTCTGAGACCTTGCAGCAGTTGGATTTAGCGGGGCACTGT 2461  
Qy 949 ValSerSerAspGlyTrpLeuAlaPheMetGlyValPheGluAsnLeuLysGlnLeuVal 968  
Db 2462 GTGAGCAGTACGGATGCTTTTACTTTCATGAATGTGTTTGAATCTCTGAAGCAGTTAGTG 2521

QY 969 PhePheAspPheSerThrLysGluPheLeuProAspProAla-LeuValArgLysLeuSe 988  
 DB 2522 TTTTITGACTTTAGCAGTGGAGGTTCTTACCGATGAGCCACTGGTGAGGAACCTTAG 2581  
 QY 988 rGlnValLeuSerLysLeuThrPheLeuGlnGluAlaArgLeuValGlyTrpGlnPheAs 1008  
 DB 2582 TCAAGTGTATCCAAAGTTAACTCTTCTGCAAGAGGTAAAGCTCACGGGCTGGGAGTTTGA 2641  
 QY 1008 pAspAspLeuSerValIleThrGlyAlaPheLysLeuValThrAla 1024  
 DB 2642 TGACTATGATATTAGCGCTATTAAAGGCACCTTTAAACTAGTACTGCT 2690

RESULT 2  
 AQ309404/C AQ309404 552 bp DNA linear GSS 22-DEC-1998  
 LOCUS CITBI-E1-2528J13.TF CITBI-E1 Homo sapiens genomic clone 2528J13,  
 DEFINITION genomic survey sequence.

ACCESSION AQ309404  
 VERSION AQ309404.1 GI:4041438  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 552)  
 Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,  
 Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and  
 Venter,J.C.

TITLE Use of a random human BAC End Sequence Database for Sequence-Ready  
 JOURNAL Map Building  
 COMMENT Unpublished

Other GSSs: CITBI-E1-2528J13.TF  
 Contact: Shaying Zhao, William Nierman, Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850  
 Tel: 301 838 0200  
 Fax: 301 838 0208

Email: hbs@tigr.org  
 Clones are available from Research Genetics (info@resgen.com). BAC  
 end search page:  
 http://www.tigr.org/tcdb/hungen/bac\_end\_search/bac\_end\_search.html.

Seq primer: M13-21  
 Class: BAC ends.

BASE COUNT 144 a 121 c 130 g 157 t  
 ORIGIN

FEATURES  
 source location/Qualifiers

1..552  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="2528J13"  
 /sex="male"  
 /cell\_type="sperm"  
 /clone\_lib="CITBI-E1"  
 /note="Vector: pBelOBAC11; Site\_1: EcoRI; Site\_2: EcoRI;  
 Caltech Human BAC Library D"

Alignment Scores:  
 Pred. No.: 2.25e-96 Length: 552  
 Score: 952.00 Matches: 183

Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 18.04% Indels: 0  
 DB: 28 Gaps: 0

US-09-864-921-97 (1-1024) x AQ309404 (1-552)

QY 477 MetValSerIleSerAspIleThrSerThrTyrrSerSerLeuLeuArgTyrrThrCysGly 496  
 DB 550 ATGGTTTCCATTTCGACATTACATCCATTATAGCAGCTCTCGGTACACCTGTGGG 491

QY 497 SerSerValGluAlaThrArgAlaValMetLysHisLeuAlaValTyrrGlnHisGly 516

DB 490 TCATCTGTGGAGCCACCAGGCGTGTATGAAGCACCTCGAGCAGTGTATCAACACGGC 431  
 QY 517 CysLeuLeuGlyLeuSerIleAlaLysArgProLeuTrpArgGlnGlnSerLeuGlnSer 536  
 DB 430 TGCCTTCTCGGACTTTCCATCGCAAGAGGCGCTCTCTGGAGACAGGAATCTTTGCAAAAGT 371  
 QY 537 ValLysSerThrThrGluGlnGluIleLeuLysAlaIleAenIleAenSerPheValGlu 556  
 DB 370 GTGAAAACACCACCTGAGCAAGAAATCTGAAGGCCATAAATCAATCTCTTTGTAGAG 311  
 QY 557 CysGlyIleHisLeuTyrrGlnGluSerThrSerLysSerAlaLeuSerGlnGluPheGlu 576  
 DB 310 TGTGCATCCATTATATCAAGAGATACATCCAAATCAGCCCTGAGCCAGAAATTTGAA 251  
 QY 577 AlaPhePheGlnGlyLysSerLeuTyrrIleAenSerGlyAenIleProAspTyrrLeuPhe 596  
 DB 250 GCTTTCTTTCAAGGTAAAGCTTATATATCACTCACTCAGGGAACATCCCCGATTTACTTTT 191  
 QY 597 AspPhePheGluHisLeuProAsnCysAlaSerAlaLeuAspPheIleLeuAspPhe 616  
 DB 190 GACTTCTTTGAACATTTGCCCCAATTGTGCAAGTGTCTGTGACTTCAATTAACCTGGACTTT 131  
 QY 617 TyrGlyGlyAlaMetAlaSerTrpGluLysAlaAlaGluAspThrGlyGlyIleHisMet 636  
 DB 130 TATGGGGAGGATATGGCTTCATGGGAAAGGCTGCAGAGACACAGGTGGATCCACATG 71  
 QY 637 GluGluAlaProGluThrTyrrIleProSerArgAlaValSerLeuPhePheAenTrpLys 656  
 DB 70 GAAGAGGCCCCAGAAACCTACATTTCCCAGCAGGGCTGTATCTTTGTTCTTCAACTGGAAG 11  
 QY 657 GlnGluPhe 659  
 DB 10 CAGGAATTC 2

RESULT 3  
 BH293386 697 bp DNA linear GSS 30-NOV-2001  
 LOCUS BH293386  
 DEFINITION CH230-44G15.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone

CH230-44G15, genomic survey sequence.  
 ACCESSION BH293386  
 VERSION BH293386.1 GI:17205794  
 KEYWORDS GSS.

SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

1 (bases 1 to 697)  
 Zhao,S., Shetty,J., Shatsman,S., Tsagaye,G., Geer,K., Shvartsbeyn  
 ,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de  
 Jong,P. and Frazer,C.M.

Rat BAC End Sequences from Library CHORI-230 EcoRI segment  
 Unpublished

Other GSSs: CH230-44G15.TJ  
 Contact: Shaying Zhao

Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208

Email: szhao@tigr.org  
 Clones are derived from the rat BAC library CHORI-230  
 (http://www.chori.org/bacpac/rat230.htm). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).  
 Clones may be purchased from BACPAC Resources  
 (http://www.chori.org/bacpac/oring information.html). BAC end

page: http://www.tigr.org/tcdb/bac\_ends/rat/bac\_end\_intro.html  
 Plate: 44 row: G column: 15  
 Seq primer: T7  
 Class: BAC ends.

FEATURES  
 Location/Qualifiers  
 1..697

```

/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SsNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-44G15"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 1"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
Pieter de Jong"
BASE COUNT 182 a 165 c 190 g 160 t
ORIGIN

Alignment Scores:
Pred. No.: 3,51e-94 Length: 697
Score: 934.00 Matches: 184
Percent Similarity: 87.77% Conservative: 17
Best Local Similarity: 80.35% Mismatches: 28
Query Match: 17.70% Indels: 1
DB: 28 Gaps: 0

US-09-864-921-97 (1-1024) x BH293386 (1-697)
QY 266 LysGluAsnHisArgPheLysAsnMetValIleValThrThrThrGluCysLeuArg 285
DB 11 AAGGAAACCATCG-TTTAAGAACATGGTCAATGTGCACCAACCAACGAGTGCCTGAGG 69
QY 286 HisIleArgGlnPheGlyAlaLeuThrAlaGluValGlyAspMetThrGluAspSerAla 305
DB 70 CACATCAGACACGTTGGCGCCCTGACTGTGGAGTGGGAGATGATGACCAACGAGCGCC 129
QY 306 GlnAlaLeuIleArgGluValIleLysGluLeuAlaGluCysLeuGlnIle 325
DB 130 CGAGTTCTCATCGGAAGTGTGATAAATGAACCTGGCTGAAGGCTGTGTGTCCAGATG 189
QY 326 GlnLysSerArgCysLeuArgAsnLeuMetLysThrProLeuPheValIleThrCys 345
DB 190 CAGGATCCAGGTGTTGAGAAATCTGATGAGGACCCCTCTCTCTGGTGTGTAACCTGT 249
QY 346 AlaIleGlnMetGlyGluSerGluPheHisSerHisThrGlnThrThrLeuPheHisThr 365
DB 250 GCCATCCAGATGGCAGTGGAGAAATCCAAAGCTCACACTCAAAACCATGCTCTTCCAAAC 309
QY 366 PheTyrAspLeuLeuGlnLysAsnLysHisLysHisGlyValAlaAlaSerAsp 385
DB 310 TTCTAGACCTCTGTATACAGAAAACAGCGCCAGACACAGTGGAGAACTTCAGGTGAT 369
QY 386 PheIleArgSerLeuAspHisArgGlyAspLeuAlaLeuGluGlyValPheSerHisLys 405
DB 370 TTGTGAGGAGCTAGACTACTGTGGAGACCTGGCCCTGGAGGTGTGTCTCCACAG 429
QY 406 PheAspPheGluLeuGlnAspValSerSerValAsnGluAspValLeuLeuThrThrGly 425
DB 430 TTGTATTCCAACTTCAGCATGTGTGCAGCATGAATGAGGACGTCCTGGTGGAGCGGG 489
QY 426 LeuLeuCysLysTyThrAlaGlnArgPheLysProLysTyrlsPheHisLysSer 445
DB 490 CTCCTCTGAAGTACACGCTCAGAGGCTGAGGCGCCACGCTATATAATCTTTCAATAATCC 549
QY 446 PheGlnGluTyThrAlaGlyArgArgLeuSerSerLeuThrThrSerHisGluProGlu 465
DB 550 TTTCAGGAGTACACAGCAGTGGAGACTCAGCATGTTTGTGTGAGTCCAGAGGCGAGAG 609
QY 466 GluValThrLysGlyAsnGlyTyrlsLeuGlnLysMetValSerIleSerAspIleThrSer 485
DB 610 GAGGTGAGCAAGGGGAATAGTACTTGAAGAAAATGCTTCCATCTCCGACATCATCCTCC 669
QY 486 ThrTyrSerSerLeuLeuArgTyThr 494
DB 670 CTGTATGGTAATCTGCTCTGTACACT 696

```

RESULT 4

```

BB627584
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

```

BB627584 650 bp mRNA linear EST 26-OCT-2001
BB627584 RIKEN full-length enriched, adult male urinary bladder Mus
musculus CDNA clone 9530011P19 5', mRNA sequence.
BB627584
BB627584.1 GI:16465218
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 650)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,B.,
Wataniki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.

```

```

FEATURES
source

```

```

Location/Qualifiers
1..650
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="9530011P19"
/sex="male"
/tissue_type="urinary bladder"
/dev_stages="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male urinary
bladder"
/note="Site 1: Sall; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was

```

primed with a primer [5',  
GAGAGAGAGAGATCCAGAGACTCTTTTITTTTTTNN 3'], cDNA was  
prepared by using trehalose thermo-activated reverse  
transcriptase and subsequently enriched for full-length by  
cap-trapper. cDNA went through one round of normalization  
to Rot = 20.0 and subtraction to Rot = 370.0. Second  
strand cDNA was prepared with the primer adapter of  
sequence [5' GAGAGAGATTCGAGTTAATTAATATATCCCCCCCC  
3']. cDNA was cleaved with XhoI and BamHI. Vector: a  
modified pBluescript KS(+) after bulk excision from  
Lambda FLC I."

BASE COUNT 164 a 174 c 173 g 139 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2,086-83 Length: 650  
Score: 838.00 Matches: 164  
Percent Similarity: 83.33% Conservative: 16  
Best Local Similarity: 75.93% Mismatches: 36  
Query Match: 15.88% Indels: 0  
DB: 10 Gaps: 0

US-09-864-921-97 (1-1024) x BB627584 (1-650)

QY 129 AsnLeuLysSerThrPheThrGluProIleLeuTTPArgLysAspGlnHisHisArg 148  
DB 2 AATCTGGAGAAACCTTCACAGAACTATCATGTGAGAGAGACCATCGTCATCACCCT 61  
QY 149 ValGluGlnLeuThrLeuAsnGlyLeuLeuGlnAlaLeuGlnSerProCysIleIleGlu 168  
DB 62 GTGGAGCAGCTGACTTGGCAGCCTGCTCGAGGCTCTGAAGAGCCCTGCCTGATTGAA 121  
QY 169 GlyGluSerGlyGlyGlySerThrLeuLeuGlnArgIleAlaMetLeuTrpGlySer 188  
DB 122 GCGGAGTCTGGCAAGGGAAGTCCACCTGCTGCAGAGAAATCGTATGCTCTGGCCCT 181  
QY 189 GlyLysCysAlaLeuThrLysPheLysPheValPhePheLeuArgLeuSerArgAla 208  
DB 182 GGGGCTCGAGGCTCTGAGGGTTCAGATTAGTCTTCTTCATCCTGAGAGCGCC 241  
QY 209 GlnGlyGlyLeuPheGluThrLeuCysAspGlnLeuLeuAspIleProGlyThrIleArg 228  
DB 242 ATGGGGGACTATTCCGAACACTGTACGATCAGCTCCTGAACATACCCGACTTCATCAGC 301  
QY 229 LysGlnThrPheMetAlaMetLeuLysLeuArgGlnArgValLeuPheLeuLeuAsp 248  
DB 302 AAGCCGACCTTCAAGGCTCTGCTGCTGAAGCTACACAAGAGGAGTCTCTTCTTCGAT 361  
QY 249 GlyTyrAsnGluPheLysProGlnAsnCysProGluIleGluAlaLeuIleLysGluAsn 268  
DB 362 GGTTCAAATGAATTCATCCCAAGACTGCCAGAAATTTGAAGCCCTGATAAGGAAAC 421  
QY 269 HisArgPheLysAsnMetValIleValThrThrThrThrGluCysLeuArgHisIleArg 288  
DB 422 CATCGCTTCAAGAAATGCTGCTCACCACCAACGAGAGTCTGAGGATATCAGA 481  
QY 289 GlnPheGlyAlaLeuThrAlaGluValGlyAspMetThrGluAspSerAlaGlnAlaLeu 308  
DB 482 CATGTTGGCCCTGACTCGGAGGTGGGAGATATGACCGAAGACAGTGCCTCAAGATCTC 541  
QY 309 IleArgGluValLeuIleLysGluLeuAlaGluGlyLeuLeuLeuGlnIleGlnLysSer 328  
DB 542 ATCGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601  
QY 329 ArgCysLeuArgAsnLeuMetLysThrProLeuPheValIleThr 344  
DB 602 AGGTGCTGAGAAATCTGATGAGACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 649

RESULT 5  
BI824482  
LOCUS 741 bp mRNA linear EST 04-OCT-2001  
DEFINITION 603038854F1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5179909 5',  
mRNA sequence.

ACCESSION BI824482  
VERSION BI824482.1 GI:15936032  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 741)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LHAM1148 row: k column: 14  
High quality sequence start: 3  
High quality sequence stop: 705.

FEATURES  
source

1. 741  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="IMAGE:5179909"  
/lab\_host="DH10B"  
/clone\_lib="NIH MGC 115"  
/note="Organ: pooled brain, lung, testis; Vector:  
pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA  
source anonymous pool of 6 male brains, age range 23-27; 1  
male lung, age 27; and 1 male testis, age 69. Library is  
oligo-dr primed and directionally cloned (EcoRV site is  
destroyed upon cloning). Average insert size 1.8 kb,  
insert size range 1-3 kb. Library is normalized and  
enriched for full-length clones and was constructed by C.  
Gruber (Invitrogen). Research Genetics tracking code  
021. Note: this is a NIH MGC Library."  
BASE COUNT 233 a 147 c 164 g 197 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1,01e-80 Length: 741  
Score: 815.00 Matches: 167  
Percent Similarity: 96.55% Conservative: 1  
Best Local Similarity: 95.98% Mismatches: 1  
Query Match: 15.44% Indels: 6  
DB: 12 Gaps: 0

US-09-864-921-97 (1-1024) x BI824482 (1-741)

QY 1 MetAsnPheIleLysAspAsnSerArg--AlaLeuIleGlnArgMetGlyMetThrValI 20  
DB 228 ATGAATTTTCATAAAGGACATATAGCCGAGCCCTTATTCAAAGAAATGGGAATGACTGTTA 287  
QY 20 leLysGlnIleThrAspAspLeuPheValTrpAsnValLeuAsnArgGluValAsnI 40  
DB 288 TAAAGCAATACACAGATGACCTATTGTATGGAATGTTCTGAATCGCGAAGATAACA 347  
QY 40 leIleCysCysGluLysValGluGlnAspAlaIaArgGlyIleIleHisMetIleLeuL 60  
DB 348 TCATTGTCTGCGAGAAGGTGGAGCAGGATGCTGTAGAGGATCATTCACATGATTTTGA 407  
QY 60 ysLysGlySerGluSerCysAsnLeuPheLeuLysSerLeuLysGluTrpAsnTrpL 80  
DB 408 AAAAGGGTTTCAGAGTCTGTAAACCTCTTTCTTAATCCCTTAAAGAGTGGAACTATCTC 467  
QY 80 euPheGlnAspLeuAsnGlyGlnSerLeuPheHisGlnThrSerGluGlyAspLeuAspA 100  
DB 468 TATTTCCAGGACTTCATGGACAAAGTCTTTTTCATCAGACATCAGAGGACATTTGGAGC 527

```

QY 100 sLeuAlaGlnAspLeu-LysAspLeuTyHisThrProSerPheLeuAsnPheTyPro 119
Db 528 ATTGGCTCAGGATTACAGGACITGTACCATCCCACTCTTCTGAACTTTATCCC 587
QY 120 LeuGlyGluAspIleAspIlelePheAsnLeuLysSerThrPheThrGluProIle-Le 139
Db 588 CTTGGTGAAGATATTGACATTATTTTAACTTTGAAAGCACCTTCACAGAACCTGTGCTT 647
QY 139 utrPargLysAspGlnHisHieArgValGluGlnLeuThrLeuAsn-GlyLeuLeuG 159
Db 648 GTGGAGGAAGGACCAACCATCAGCGCTGGAGCAGCTGACCTGAATGGGTCTCCTGC 707
QY 159 InAlaLeuGlnSerProCysIleIleLeuGly 169
Db 708 AGCTCTCCAGAG-CCTGCATCATTTGAAGG 738

RESULT 6
AV719179/c
LOCUS AV719179 GUC Homo sapiens cDNA clone GLCEQA10 5', mRNA sequence.
DEFINITION AV719179
ACCESSION AV719179
VERSION AV719179.1 GI:10816331
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
MAMMALIA; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 480)
AUTHORS Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H.,
XU,X., LI,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Zeng,L.,
Y.,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang
Y., Gu,Y., Chen,Z. and Han,Z.
Homo sapiens cDNA GLC clones
Unpublished
JOURNAL Chinese National Human Genome Center at Shanghai
COMMENT 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
Location/Qualifiers
1..480
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GLCEQA10"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GLC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 138 a 120 c 85 g 137 t
ORIGIN

Alignment Scores:
Pred. No.: 1..896-79 Length: 480
Score: 801.00 Matches: 156
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.18% Indels: 0
DB: 9 Gaps: 0

US-09-864-921-97 (1-1024) x AV719179 (1-480)

QY 869 HisGluLeuIleAspArgMetAsnValLeuGluGlnLeuThrAlaLeuMetLeuProTrp 888
Db 480 CATGAAGTATGACAGAGGATGAACGTGTAGAACAAGCTCAGCGCACTGATGTCGCCCTGG 421
QY 889 GlyCysAspValGlnGlySerLeuSerLeuLysHisLeuGluValProGln 908

```

```

Db 420 GGCTGTGACGTGCAAGGCAGCCTGAGCAGCCTGTTGAAACATTTGGAGGAGGTCCACAA 361
QY 909 LeuValLysLeuGlyLeuLysAsnTrpArgLeuThrAspThrGluIleArgIleLeuGly 928
Db 360 CTCGTCAAGCTTGGGTGAAAAAAGCTGGAGACTCAGATACAGATAGAGATTTAGGT 301
QY 929 AlaPhePheGlyLysAsnProLeuLysAsnPheGlnGlnLeuAsnLeuAlaGlyAsnArg 948
Db 300 GCATTTTGGAAAGAACCCCTCTGAAAAAATTCAGCAGATTGAATTTGGCGGGAATCGT 241
QY 949 ValSerSerAspGlyTrpLeuAlaPheMetGlyValPheGluAsnLeuLysGlnLeuVal 968
Db 240 GTGAGCAGTGATGATGGCTTGCCTTCATGGGTGATTTCAGAAATCTTAAAGCAATTAGTG 181
QY 969 PhePheAspPheSerThrLysGluPheLeuProAspProAlaLeuValArgLysLeuSer 988
Db 180 TTTTGTGACTTTAGTACTAAAGAATTTCTCCTGATCCAGCATTAGTCAGAAAACTTAGC 121
QY 989 GlnValLeuSerLysLeuThrPheLeuGlnGlnAlaArgLeuValGlyTrpGlnPheAsp 1008
Db 120 CAAGTGTATCAAGTTAACTTTCTGCAAGAGCTAGGCTTGTGGGTGGCAATTTGAT 61
QY 1009 AspAspLeuSerValIleThrGlyAlaPheLysLeuValThrAla 1024
Db 60 GATGATGATCTCAGTGTATTACAGGAGCTTTTAACTAGTAACTGCT 13

RESULT 7
AV360053/c
LOCUS AV360053 602 bp DNA linear GSS 02-OCT-2000
DEFINITION IM0103H11F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0103H11 F, genomic survey sequence.
ACCESSION AV360053
VERSION AV360053.1 GI:10473753
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 602)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished
JOURNAL Contact: Robert B. Weiss
COMMENT University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0103 row: H column: 11
Seq primer: CGTTGTAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 602.
Location/Qualifiers
1..602
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0103H11"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource

```

(<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 139 a 153 c 139 g 171 t

ORIGIN

Alignment Scores:  
Pred. No.: 4,598-79 Length: 602  
Score: 799.00 Matches: 154  
Percent Similarity: 88.21% Conservative: 18  
Best Local Similarity: 78.97% Mismatches: 23  
Query Match: 15.14% Indels: 0  
DB: 28 Gaps: 0

US-09-864-921-97 (1-1024) x AZ360053 (1-602)

Qy 382 AlaAlaSerAspPheIleArgSerLeuAspHisArgGlyAspLeuAlaLeuGluGlyVal 401  
Db GCTTCAGGTGATTTGCCAGGAGCTAGACTCTGTGGAGACCTGGCCCTAGAGGTGTG 528  
Qy 402 PheSerHisLysPheAspPheGluLeuGlnAspValSerSerValAsnGluAspValLeu 421  
Db TTGCGCCCAAAATTTGATTTTGAACCCGAGCATGGTCCAGCATGAACGAGGACGCTCTG 468  
Qy 422 LeuThrThrGlyLeuLeuCysLysThrAlaGlnArgPheLysProLysThrLysPhe 441  
Db GTGCAATAGGGCTCTCTGTAAAGTACACAGCTCAGAGGCTGAAGCCACCGTATAAATTC 408  
Qy 442 PheHisLysSerPheGluThrAlaGlyArgArgLeuSerSerLeuThrSer 461  
Db TTTCATAATCATTTTCAGAGGTACACGCGAGTCCGAGACTCAGCAGTTTCTGACGTC 348  
Qy 462 HisGluProGluGluValThrLysGlyAsnGlyThrLeuGlnLysMetValSerLysSer 481  
Db AAAGAGCCAGAGGAGGTGAGCAAGGAAACAGCTACTTAAACAAATGGTTTCCATCTCT 288  
Qy 482 AspIleThrSerThrTyrrSerSerLeuLeuArgThrCysGlySerSerValGluAla 501  
Db GACATCATCATCCCTATATGGCAATCTGCTCTACAGTGTGGTGTGTCACAGAGCA 228  
Qy 502 ThrArgAlaValMetLysHisLeuAlaValTyrrGlnHisGlyCysLeuLeuGlyLeu 521  
Db ACCAGGGCGGTGATGAGGACCTTGCATGTTTATCAGCAGGAGGCTTACAGGACTT 168  
Qy 522 SerIleAlaLysArgProLeuThrPargGlnGlnSerLeuGlnSerValLysAsnThrThr 541  
Db TCAGTCACCAAGAGGCTCTCTGGAGGAGGAAATCAATCCAGAGTCTGAGAAATACCACT 108  
Qy 542 GluGlnGluIleLeuLysAlaIleAsnLysLeuSerPheValGluCysGlyLysLeu 561  
Db GAGCAAGATGTTCTGAAGGCCATCAATGTAAATCCCTCGTAGAGTGGGCATCAATTTG 48  
Qy 562 TyrGlnGluSerThrSerLysSerAlaLeuSerGlnGluPheGlu 576  
Db TTCTCAGAGAGTATGCTTAATCAGACCTGAGCCAGAAATTGAA 3

RESULT 8

LOCUS BI908869

DEFINITION 603066455F1 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5215669 5', 748 bp mRNA linear EST 16-OCT-2001

mRNA sequence.

ACCESSION BI908869  
VERSION BI908869.1 GI:16171950

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 748)

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: [rsb@nih.gov](mailto:rsb@nih.gov)

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LUAM1541 row: m column: 14

High quality sequence start: 7

High quality sequence stop: 744.

FEATURES

Location/Qualifiers

1..748

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5215669"

/tissue\_type="leukocyte"

/lab\_host="DH10B"

/clone\_lib="NIH\_MGC\_118"

/notes="vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH MGC Library."

BASE COUNT 236 a 148 c 166 g 198 t

ORIGIN

Alignment Scores:

Pred. No.: 1,97e-76 Length: 748  
Score: 777.00 Matches: 159  
Percent Similarity: 95.81% Conservative: 1  
Best Local Similarity: 95.21% Mismatches: 5  
Query Match: 14.72% Indels: 4  
DB: 12 Gaps: 0

US-09-864-921-97 (1-1024) x BI908869 (1-748)

Qy 1 MetAsnPheIleLysAspAsnSerArgAlaLeuIleGlnArgMetGlyMetThrValIle 20  
Db ATGAATTTCTATAAGGACAAATAGCCGAGCCCTTATTCAAAGAAATGGGAATGACTGTATA 311  
Qy 21 LysGlnIleThrAspAspLeuPheValTrpAsnValLeuAsnArgGluGluValAsnIle 40  
Db AAGCAATATCAGATGACCTATTGTGTATGGATGTTCTGAATCCGAGGAAGTAACATC 371  
Qy 41 IleCysCysGluLysValGluInAspAlaAlaArgGlyIleIleHisMetIleLeuLys 60  
Db ATTTGCTGCGAAGAGTGGAGCAGATGCTGTAGAGGATCATTCATCATGATTTGAAA 431  
Qy 61 LysGlySerGluSerCysAsnLeuPheIleLysSerLeuLysGluTrpAsnTyrProLeu 80  
Db AAGGTTTCAGAGTCTGTAACTCTTTCTTAAATCCCTTAAGGAGTGAATCTCTCTCTA 491  
Qy 81 PheGlnAspLeuAsnGlyGlnSerLeuPheHisGlnThrSerGluGlyAspLeuAspAap 100



Db 492 TTTCCAGGACTTGAATGGCAAAAGCTCTTTTTCATCAGACATCAGAGGAGACTTGGACGAT 551  
 QY 101 LeuAlaGluAspLeu--LysAspLeuTyHisThrProSerPheLeuAsnPheTyProL 120  
 Db 552 TTGGCTCAGATTTCAGAGGACTTGTACCATACCCATCTTTTCTGAACTTTTATCCCC 611  
 QY 120 euGlyGluAspLeuAspLeuLeuPheAsnLeuLysSerThrPheThrGluProIleLeu 140  
 Db 612 TTGGTCAAGATATTGACATATTATTTTAACCTTGAAGAGCCTTTCACAGAGCTGTCTGT 671  
 QY 140 rpArgLysAspGlnHisHisArgValGluGlnLeuThrLeuAsnGlyLeuLeuGlnA 160  
 Db 672 GGAGGAGGAGCAAAAC-CATCACCGGTGAGCAGCTGACCTGAATGGCTC-CTGCAGG 729  
 QY 160 laLeuGlnSerProCys 165  
 Db 730 CTCTTCAGAGCCCTGTC 746

## RESULT 9

BH348412 518 bp DNA linear GSS 03-DEC-2001  
 LOCUS CH230-42F7.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone  
 DEFINITION CH230-42F7, genomic survey sequence.

## ACCESSION

BH348412

## VERSION

BH348412.1 GI:17279146

## KEYWORDS

GSS.

## SOURCE

Rattus norvegicus (Norway rat)

## ORGANISM

Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

## REFERENCE

1 (bases 1 to 518)  
 Zhao,S., Shetty,J., Shatsman,S., Tsengay,G., Geer,K., Shvartsbeyn  
 ,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de  
 Jong,P. and Fraser,C.M.

## TITLE

Rat BAC End Sequences from Library CHORI-230 EcoRI segment

## JOURNAL

Unpublished

## COMMENT

Other GSSs: CH230-42F7.TJ  
 Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org  
 Clones are derived from the rat BAC library CHORI-230  
 (http://www.chori.org/bacpac/rat230.htm). For BAC library  
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).  
 Clones may be purchased from BACPAC Resources  
 (http://www.chori.org/bacpac/or ering information.htm). BAC end  
 page: http://www.tigr.org/tdb/bac\_ends/rat/bac\_end\_intro.html  
 Plate: 42 row: F column: 7  
 Seq primer: T7  
 Class: BAC ends.

## FEATURES

## source

## Location/Qualifiers

1..518  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /strain="BN/SsNHsd/MCW"  
 /db\_xref="taxon:10116"  
 /clone="CH230-42F7"  
 /sex="Female"  
 /cell\_type="Brain"  
 /clone\_lib="CHORI-230 Segment 1"  
 /notes="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;  
 CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by  
 Pieter de Jong"

## BASE COUNT

133 a 124 c 146 g 115 t

## ORIGIN

## Alignment Scores:

Pred. No.: 3.2e-67 Length: 518

Score: 693.00 Matches: 136

Percent Similarity: 86.71% Conservative: 14  
 Best Local Similarity: 78.61% Mismatches: 22  
 Query Match: 13.13% Indels: 1  
 DB: 28 Gaps: 0

US-09-864-921-97 (1-1024) x BH348412 (1-518)

QY 261 lIleGluAlaLeuIleLysGluAsnHisArgPheLysAsnMetValIleValThrThr 280  
 Db 2 ATCGAGGCCCTGGTAAGGAAAACCATCGTTTCAGACATCGTTCATTGTGCACCACC 61  
 QY 281 ThrGluCysLeuArgHisIleArgGlnPheGlyAlaLeuThrAlaGluValGlyAspMet 300  
 Db 62 ACGGAGTGCCTCAGGCACATCAGACACAGCTGTGGCGCCTGACTGTGGAGGTGGAGATG 121  
 QY 301 ThrGluAspSerAlaGlnAlaLeuIleArgGluValLeuIleLysGluLeuAlaGluGly 320  
 Db 122 ACCGAAGACAGCGCCCGAGTCTTCATCCGGGAAGTGTCTGATAAATGAACCTGGCTGAAGC 181  
 QY 321 LeuLeuLeuGlnIleGlnLysSerArgCysLeuArgAsnLeuMetLysThrProLeuPhe 340  
 Db 182 TTGTTGTTCCAGATCGAGAGTCCAGGTGCTTGAGAAATCTGATGAGGACCCCTCTCTTC 241  
 QY 341 ValValIleThrCys-AlaIleGlnMetGlyGluSerGluPheHisSerHisThrGlnTh 360  
 Db 242 GTGGTGATAACCTGTGACCATCCAGATGGCAGTGAGGAATTCGAAGCTCACACTCAAAC 301  
 QY 360 rThrLeuPheHisThrPheTyAspLeuIleGlnLysAsnLysHisLysHisLysG 380  
 Db 302 CATGCTCTTCCAAACCTTCTACGACCTCTCTGATACAGAAAAACAGCGCGACACAGTGG 361  
 QY 380 vValAlaAlaSerAspPheIleArgSerLeuAspHisArgGlyAspLeuAlaLeuGlu 400  
 Db 362 AGGAATCTCAGTGATTTTGTTCAGGAGCCTAGACTACTGTGGAGACCTGGCCCTGAAGG 421  
 QY 400 vValPheSerHisLysPheAspPheGluLeuGlnAspValSerSerValAsnGluAspVa 420  
 Db 422 TGTGTTCTCCCAAGTTTGATTTCGAACTTGAGGATGTGTGCAGCATGATGAGGACCT 481  
 QY 420 lleuLeuThrThrGlyLeuLeuCysLysTyThrAla 432  
 Db 482 CCTGGTGAGGACGGGGCTCTCTGTGTAAGTACAGCG 518

## RESULT 10

BH267158/c

## LOCUS

BH267158

## DEFINITION

CH230-19B22.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone

CH230-19B22, genomic survey sequence.

## ACCESSION

BH267158

## VERSION

BH267158.1 GI:17179098

## KEYWORDS

GSS.

## SOURCE

Rattus norvegicus (Norway rat)

## ORGANISM

Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

## REFERENCE

1 (bases 1 to 619)

## AUTHORS

Zhao,S., Shetty,J., Shatsman,S., Tsengay,G., Geer,K., Shvartsbeyn  
 ,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de  
 Jong,P. and Fraser,C.M.

## TITLE

Rat BAC End Sequences from Library CHORI-230 EcoRI segment

## JOURNAL

Unpublished

## COMMENT

Other GSSs: CH230-19B22.TV  
 Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org  
 Clones are derived from the rat BAC library CHORI-230  
 (http://www.chori.org/bacpac/rat230.htm). For BAC library  
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources  
 (http://www.choxi.org/bacpac/or ering information.html). BAC end  
 page: http://www.tigr.org/cdb/bac\_ends/rat/bac\_end\_intro.html  
 Plate: 19 row: B column: 22  
 Seq primer: SP6  
 Class: BAC ends.

# FEATURES

## Location/Qualifiers

1..619  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /strain="BN/SsHsd/MCW"  
 /db\_xref="taxon:10116"  
 /clone="CH230-19B22"  
 /sex="Female"  
 /cell\_type="Brain"  
 /clone\_lib="CHORI-230 Segment 1"  
 /note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;  
 CHORI-230 Rat (BN/SsHsd/MCW) BAC library produced by  
 Pieter de Jong"

BASE COUNT 163 a 141 c 158 g 157 t

## Alignment Scores:

Pred. No.: 3,12e-57 Length: 619  
 Score: 605.50 Matches: 119  
 Percent Similarity: 80.34% Conservative: 24  
 Best Local Similarity: 66.85% Mismatches: 30  
 Query Match: 11.47% Indels: 5  
 DB: 28 Gaps: 2

US-09-864-921-97 (1-1024) x BH267158 (1-619)

QY 73 LeuLysGluThrPAsnTyrProLeuPheGln-----AspLeuAsnGlyGlnSerLeu 89  
 DB 530 CTTCAGATGCAATCATCCCTCTATACATCATATATCTCTAACAGGT-----CTT 477  
 QY 90 PheHisGlnThrSerGluGlyAspLeuAspLeuAlaGlnAspLeuLysAspLeuTyr 109  
 DB 476 TTTCATCAGAACTTGAAGAAGACTTGGATGTTCTGGCCAGAGTCTAAAGACTTATAC 417  
 QY 110 HisThrProSerPheLeuAsnPheTyrProLeuGlyGluAspIleAspIleIlePheAsn 129  
 DB 416 AACAGCCCTGTTTAAAGAACTTCTTCTCTGGGTGAGGATATCGACATCATTTTCAAT 357  
 QY 130 LeuLysSerThrPheThrGluProIleLeuTyrArgLysAspGlnHisHisArgVal 149  
 DB 356 CTGCAGATCACCTTCACAGACCTGTCTTGTGGAGGAGGACCATCGTCATCCCGAGTA 297  
 QY 150 GluGlnLeuThrLeuAsnGlyLeuLeuGlnAlaLeuGlnSerProCysIleIleGluGly 169  
 DB 296 GAGCAGATGACCTTGGCGAGCTGTGGAGGCTCTGAAGAGTCCCTGCTCATTAAGAGG 237  
 QY 170 GluSerGlyLysGlyLysSerThrLeuLeuGlnArgIleAlaMetLeuTrpGlySerGly 189  
 DB 236 GAGTCTGGCAAGGGAAGTCCACCTCTGTCGCAAAAAATTCATGCTCTGGCCCTATGGA 177  
 QY 190 LysCysLysAlaLeuThrLysPheLysPheValPhePheLeuArgLeuSerArgAlaGln 209  
 DB 176 ATGTGCCAGCTCTGAACAGTTCAAATTTGGTCTTTTCATCCGCTGAGCAGTGTAGG 117  
 QY 210 GlyGlyLeuPheGluThrLeuCysAspGlnLeuLeuAspIleProGlyThrIleArgLys 229  
 DB 116 GTGGGCTCTTTGAACATGTATGATCAGCTCGTGAACATACCTGACTCCATCAGCAA 57  
 QY 230 GlnThrPheMetAlaMetLeuLysLeuArgGlnArgValLeuPheLeuLeu 247  
 DB 56 CCAACCTTCAGGCTCTGCTGCTGAAGTTTACACAAGAAAGTCTCTTCTCTCTC 3

## RESULT 11

AV656315 371 bp mRNA linear EST 16-JAN-2002  
 LOCUS  
 DEFINITION AV656315 GLC Homo sapiens cDNA clone GLCEQA10 3', mRNA sequence.  
 ACCESSION AV656315

AV656315.1 GI:9877329

EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,  
 Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,  
 Shen,K., Lu,G., Fu,G., Zhong,W., Xu,S., Gu,W., Huang,W., Zhao,X.,  
 Hu,G., Gu,J., Chen,Z. and Han,Z.

TITLE  
 1 (bases 1 to 371)  
 Insight into hepatocellular carcinogenesis at transcriptome level  
 by comparing gene expression profiles of hepatocellular carcinoma  
 with those of corresponding noncancerous liver

JOURNAL  
 MEDLINE  
 PUBMED  
 COMMENT Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)  
 21625106  
 11752456

Contact: Zeguang Han

Chinese National Human Genome Center at Shanghai

351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai

201203, P. R. China

Tel: 86-21-50801919 (ex.45)

Fax: 86-21-50801922

Email: hanzg@chgc.sh.cn

This clone is available at CHGC in Shanghai.

## FEATURES

### source

1..371  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="GLCEQA10"  
 /tissue\_type="corresponding non cancerous liver tissue"  
 /dev\_stages="Adult"  
 /lab\_host="SOLR"  
 /clone\_lib="GLC"  
 /note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2:  
 XhoI"

BASE COUNT 112 a 85 c 91 g 82 t 1 others

## Alignment Scores:

Pred. No.: 7,38e-57 Length: 371  
 Score: 599.00 Matches: 123  
 Percent Similarity: 87.23% Conservative: 0  
 Best Local Similarity: 87.23% Mismatches: 0  
 Query Match: 11.35% Indels: 18  
 DB: 9 Gaps: 1

US-09-864-921-97 (1-1024) x AV656315 (1-371)

QY 625 GluLysAlaAlaGluAspThrGlyGlyIleHisMetGluGluAlaProGluThrTyrIle 644  
 DB 1 GAAAGGCTGCAGAGACACAGAGTGAATCCACATGGAAGAGGCCGCCAAGAACCTACATT 60  
 QY 645 ProSerArgAlaValSerLeuPheAsnTrpLysGlnGluPheArgThrLeuGluVal 664  
 DB 61 CCCAGAGGGCTGTATCTTTCTTCACTGGAAGCAGGAATTCAGAGCTCTGGAGTTC 120  
 QY 665 ThrLeuArgAspPheSerLysLeuAsnLysGlnAspIleArgTyrLeuGlyLysIlePhe 684  
 DB 121 ACACCTCGGGATTTTCAGCAAGTTGAATAAGCAAGATATCATATCTGGGAAAAATATTC 180  
 QY 695 SerSerAlaThrSerLeuArgLeuGlnIleLysArgCysAlaGlyValAlaGlySerLeu 704  
 DB 181 AGCTTGCACAAAGCTCAGGCTCAATAAAGAGATGTCTGTGTGTGGTGGAGGAGCCTC 240  
 QY 705 SerLeuValLeuSerThrCysLysAsnIleTyrSerLeuMetValGluAlaSerProLeu 724  
 DB 241 AGTTTGTCTTCAGCACCCTGTGAACATTTATTTCTCTCATGTGGAGGAGCAGTCCCTC 300  
 QY 725 ThrIleGluAspGluArgHisIleThrSerValThrAsnLeuLysThrLeuSerIleHis 744  
 DB 301 ACCATAGAAGATGAGAGGCACATCATCTCTGTAAACAACTCTG----- 342

```

QY 745 AspLeuGlnAsnGlnArgLeuProGlyGlyLeuThrAspSerLeuGlyAsnLeuLysAsn 764
Db 343 -----GACAGCTTGGTAACTTGAAGAC 366

QY 765 Leu 765
Db 367 CTT 369

RESULT 12
AI263294/c 364 bp mRNA linear EST 03-FEB-1999
LOCUS qx57b01.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2005417 3',
DEFINITION mRNA sequence.
ACCESSION AI263294
VERSION AI263294.1 GI:3871497
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert length: 2146 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 364.
FEATURES
Location/Qualifiers
1..364
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2005417"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Pan1"
/notes="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"
BASE COUNT 117 a 84 c 55 g 108 t
ORIGIN
Alignment Scores:
Pred. No.: 2,71e-55 Length: 364
Score: 585.00 Matches: 114
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.09% Indels: 0
DB: 9 Gaps: 0

US-09-864-921-97 (1-1024) x AI263294 (1-364)

QY 911 LysLeuGlyLeuLysAsnTrpArgLeuThrAspThrGluLeuArgIleLeuGlyAlaPhe 930
Db 364 AAGCTTGGGTGGAAGAACTGGAGACTCACAGATACAGATTAGAAATTTAGTGCAATTT 305

QY 931 PheGlyLysAsnProLeuLysAsnPheGlnGlnLeuAsnLeuAlaGlyAsnArgValSer 950
Db 304 TTGGAAGAACCTCTGAAAACTTCACAGCTTGAAATTTGGCGGAATCTGTGAGC 245

QY 951 SerAspGlyTrpLeuAlaPheMetGlyValPheGluAsnLeuLysGlnLeuValPhePhe 970
Db 244 AGTGATGGATGGCTTGCCTTCATGGGTGATTGAGAACTCTTAAGCAATTAGTGTGTTTTT 185

```

```

QY 971 AspPheSerThrLysGluPheLeuProAspProAlaLeuValArgLysLeuSerGlnVal 990
Db 184 GACTTTAGTACTAAGAATTTCTACTGATCCAGATTAGTCAGAAAACCTTAGCCAAATG 125

QY 991 LeuSerLysLeuThrPheLeuGlnGlnAlaArgLeuValGlyTrpGlnPheAspAsp 1010
Db 124 TATCCAAAGTTAACTTTCTGCAAGAAGCTAGGCTTGTGGGTGGCAATTGATGATGAT 65

QY 1011 AspLeuSerValIleThrClyAlaPheLysLeuValThrAla 1024
Db 64 GATCTCAGTGTATTACAGGTCGCTTTTAAACTAGTAACGTCT 23

RESULT 13
BI854236
LOCUS BI854236
DEFINITION 603381263F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:5389239 5',
ACCESSION BI854236
VERSION BI854236.1 GI:15994983
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 775)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furch
Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLA411991 row: m column: 16
High quality sequence stop: 764.
FEATURES
Location/Qualifiers
1..775
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5389239"
/tissue_type="tumor, gross tissue"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam4"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Priscilla Furch,
NIH Reference for transgenic model: Li et al., Cell Growth
and Differentiation 7, 3-11 (1996)."
BASE COUNT 226 a 162 c 196 g 191 t
ORIGIN
Alignment Scores:
Pred. No.: 9,18e-55 Length: 775
Score: 595.00 Matches: 118
Percent Similarity: 78.66% Conservative: 11
Best Local Similarity: 71.95% Mismatches: 35
Query Match: 11.09% Indels: 0
DB: 12 Gaps: 0

US-09-864-921-97 (1-1024) x BI854236 (1-775)

QY 861 GluLysAspGlyAsnGluAlaLeuHisGluLeuIleAspArgMetAsnValLeuGln 880
Db 861 GATCTCAGTGTATTACAGGTCGCTTTTAAACTAGTAACGTCT 23

```

Db 2 GAAAGGATGGGAATGAAGCTCTACAGGAACATGATCGGCAGGCTTGGCGTCTCTGGAGAG 61

Qy 881 LeuThrAlaLeuMetLeuProTrpGlyCysAspValGlnGlySerLeuSerSerLeuLeu 900

Db 62 CTCACATACATGATGCTGCTGGTCTGGTCTGGATGTGCACACGAGCTGCCCAAGCTGTG 121

Qy 901 LysHisLeuGluGluValProGlnLeuValLysLeuGlyLeuLysAsnTrpArgLeuThr 920

Db 122 AGCAGTTGAGGGGACCCAGGACTTGCACAACTGGATTGAAAACTGGAGACTCAGA 181

Qy 921 AspThrGluLeuArgLeuGlyAlaPhePheGlyLysAsnProLeuLysAsnPheGln 940

Db 182 GACGAAGAGATTAAAGATTAGTGAATTTCTGGAGATGAATCTCTGAGAGACTTGCAG 241

Qy 941 GlnLeuAsnLeuAlaGlyAsnArgValSerSerAspGlyTrpLeuAlaPheMetGlyVal 960

Db 242 CAGTTGGATTAGCGGGGACATGTGTGACAGTGCAGGATGGCTTTACTTCAATGATGTG 301

Qy 961 PheGluAsnLeuLysGlnLeuValPhePheAspPheSerThrLysGluPheLeuProAsp 980

Db 302 TTTGAGAACTGAGCAGTTAGTGTGTTTGTGCTTTAGCACTGAGGAGTTCTTACCGAT 361

Qy 981 ProAlaLeuValArgLysLeuSerGlnValLeuSerLysLeuThrPheLeuGlnGluAla 1000

Db 362 GCAGCACTGCTGAGGAACCTTAGTCAAGTGTATCCAACTTAACCTCTGCAAGAGGTA 421

Qy 1001 ArgLeuValGlyTrpGlnPheAspAspAspLeuSerValIleThrGlyAlaPheLys 1020

Db 422 AAGCTCACGGCTGGGAGTTGATGACTATGATATTAGCGCTATTAAAGGCACCTTTAA 481

Qy 1021 LeuValThrAla 1024

Db 482 CTAGTACCTGCT 493

## RESULT 14

BH358172/c  
LOCUS  
DEFINITION  
CH230-18E7.TJB CHORI-230 Segment 1 Rattus norvegicus genomic clone  
CH230-18E7, genomic survey sequence.

## ACCESSION

BH358172

## VERSION

BH358172.1 GI:17288906

## KEYWORDS

GSS.

## SOURCE

Rattus norvegicus (Norway rat)

## ORGANISM

Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

## REFERENCE

1 (bases 1 to 817)  
Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K., Shvartabeyn  
A., Gebregorgis, E., Overton, L., Russell, D., Chen, D., Riggs, F., de  
Jong, P. and Fraser, C.M.

## TITLE

## JOURNAL

## COMMENT

Other GSSs: CH230-18E7.TVB  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org  
Clones are derived from the rat BAC library CHORI-230  
(<http://www.chori.org/bacpac/rat230.htm>). For BAC library  
availability, please contact Pieter de Jong (pdejong@mail.choi.org).  
Clones may be purchased from BACPAC Resources  
(<http://www.chori.org/bacpac/or> exing information.html). BAC end  
page: [http://www.tigr.org/tadb/bac\\_end5/rat/bac\\_end\\_intro.html](http://www.tigr.org/tadb/bac_end5/rat/bac_end_intro.html)  
Plate: 18 row: E Column: 7

## Seq primer: SP6

## Class: BAC ends.

## Location/Qualifiers

## 1. .817

## /organism="Rattus norvegicus"

## FEATURES

## source

/mol\_type="genomic DNA"  
/strain="BN/SnHsd/MCW"  
/db\_xref="taxon:10116"  
/clone="CH230-18E7"  
/sex="Female"  
/cell\_type="Brain"

/clone\_lib="CHORI-230 Segment 1"

/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: BcoRI;

CHORI-230 Rat (BN/SnHsd/MCW) BAC library produced by

Pieter de Jong"

BASE COUNT 211 a 182 c 199 g 225 t

## ORIGIN

Alignment Scores:

Pred. No.: 2,05e-52 Length: 817

Score: 564.50 Matches: 121

Percent Similarity: 74.6% Conservative: 25

Best Local Similarity: 67.36% Mismatches: 37

Query Match: 10.70% Indels: 7

DB: 28 Gaps: 2

US-09-864-921-97 (1-1024) x BH358172 (1-817)

Qy 62 GlySerGluSerCysAsnLeuPheLeuLysSerLeuLysGluTrpAsnTrpProLeuPhe 81

Db 556 GGCATCTCTGCCATTCATCTCCCAAGCCTTGGGCTTCGAGATGCAATCATGCTCTAT 497

Qy 82 Gln-----AspLeuAsnGlyGlnSerLeuPheHisGlnThrSerGluGlyAspLeu 98

Db 496 AACATCATATTTCTCTAACAGGT-----CTTTTTCATCAGAACTTGGAGAGACTTG 443

Qy 99 AspAspLeuAlaGlnAspLeuLysAspLeuTrpHisThrProSerPheLeuAsnPheTyr 118

Db 442 GATGTTCTGGGCCAGAGTCTAAAGACTTATACACAGCCCTGTTTTTAAGAACTTCITT 383

Qy 119 ProLeuGlyGluAspIleAspIlePheAsnLeuLysSerThrPheThrGluProIle 138

Db 382 CCTCTGGGTGAGGATATCGACATCATTTTCAATCTGCAGATCACCTTACAGAACTGTC 323

Qy 139 LeuTrpArgLysAspGlnHisHisArgValGluGlnLeuThrLeuAsnGlyLeuLeu 158

Db 322 TTGTGGAGGAAGGACCATCGTCATCACCGAGTAGAGCAGATGACCTGGCAGCCTGCTG 263

Qy 159 GlnAlaLeuGlnSerProCysIleIleGluGlyGluSerGlyLysGlySerThrLeu 178

Db 262 GAGGCTCTGAAGAGTCCCTGCTCATTTGAAGGGAGTCTGGCAAGGGAAGTCCACCTG 203

Qy 179 LeuGlnArgIleAlaMetLeuTrpGlySerGlyLysCysLysAlaLeuThrLysPheLys 198

Db 202 CTGCAAAAATTCATCTCTGGGCTCTGGGATGTGCCAGCTCTGAACCAAGTTCAAA 143

Qy 199 PheValPhePheLeuArgLeuSerArgAlaGlnGlyLeuPheGluThrLeuCysAsp 218

Db 142 TTGTGCTTCTTCATCCGCTGAGCAGTCTAGGGGTGGCTGTTTGAACAATTTGA-GAT 84

Qy 219 GlnLeuLeuAspIleProGlyThrIleArgLysGlnThrPheMetAlaMetLeuLys 238

Db 83 CAGCTCGTGAACATACCTGACTTCCATCAGCAAAACCACTTCAGGGCTCTGCTGCT-AAG 25

Qy 239 LeuArgGlnArgValLeuPheLeu 246

Db 24 TTACACAGAAAGTCTCTTCTC 1

## RESULT 15

BX109471/c

## LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

BX109471 503 bp mRNA linear EST 07-FEB-2003

IMAGE:1909558, mRNA sequence.

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 503)  
 AUTHORS Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M., Radelof, U., Schneider, D. and Korn, B.  
 TITLE Human Unigeneset - RZPD3  
 JOURNAL Unpublished  
 COMMENT Contact: Ina Rolfs  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
 RZPD: IMAGP998B234682.  
 RZPDLIB: I.M.A.G.E. cDNA Clone Collection;  
 Human Unigeneset - RZPD3 (RZPDLIB No.972)  
 http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Heubnerweg 6, D-14059 Berlin, Germany  
 Tel: +49 30 32639 101  
 Fax: +49 30 32639 111  
 www.rzpd.de

This clone is available royalty-free from RZPD:  
 contact RZPD (clone@rzpd.de) for further information. Seq primer:  
 M13r, Primer sequence: TTTCACAGGAACAGCTATGAC.

#### FEATURES

Location/Qualifiers  
 1..503  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGP998B234682 ; IMAGE:1909558"  
 /tissue\_type="carcinoid"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP Lu5"  
 /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from neuroendocrine lung carcinoid, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 127 a 111 c 128 g 128 t 9 others  
 ORIGIN

#### Alignment Scores:

Pred. No.: 1,38e-52 Length: 503  
 Score: 563.00 Matches: 106  
 Percent Similarity: 96.40% Conservative: 1  
 Best Local Similarity: 95.50% Mismatches: 4  
 Query Match: 10.67% Indels: 0  
 DB: 13 Gaps: 0

US-09-864-921-97 (1-1024) x EX109471 (1-503)

QY 89 LeuPheHisGlnThrSerGluGlyAspLeuAspLeuAlaGlnAspLeuLysAspLeu 108  
 Db 334 CTTTTCATCAGACATCAGAGGAGNCTTGGACGATTTCGCTCAGGATTTAAGGNCITG 275  
 QY 109 TyrHisThrProSerPheLeuLeuPheTyrProLeuGlyGluAspIleAspIlelePhe 128  
 Db 274 TNCATNCCCCTCTTTTCTGAACCTTTATCCCTTGGTGAAGATATTGACATATTTT 215  
 QY 129 AsnLeuLysSerThrPheThrGluProIleLeuTrpArgLysAspGlnHisHisArg 148  
 Db 214 AACTTGAAGAAGCACCTTCACAGAACTCTCTGTGGAGGAGGACCAACACCATCACC 155  
 QY 149 ValGluGlnLeuThrLeuLeuGlyLeuLeuGlnAlaLeuGlnSerProCysIleIleGlu 168  
 Db 154 GTGGACAGCTGACCTCGATGGCTCTCTGCAGGCTCTTCAGAGCCCTCGCATCTGAA 95  
 QY 169 GlyGluSerGlyLysGlyLysSerThrLeuLeuGlnArgIleAlaMetLeuTrpGlySer 188  
 Db 94 GGGGAATCTGGCAAGGCAAGTCCACTCTGCTGCAGCAATTGCCATGCTCTGGGGCTCC 35

QY 189 GlyLysCysLysAlaLeuThrLysPheLysPhe 199  
 Db 34 GGAAAGTCAAGGCTCTGACCAAGTTCAAATTC 2

Search completed: December 13, 2003, 04:24:54  
 Job time : 3346 secs

THIS PAGE BLANK (USPTO)